

WHAT IS CLAIMED IS:

1. A scalable clustered system, comprising:

a global fabric; and

two or more cluster nodes interconnected via the global fabric, each cluster node including a node naming agent (NNA), a local fabric and one or more end nodes interconnected via the local fabric, the NNA being configured as a fully symmetrical translation device interposed between the local fabric and the global fabric, the NNA providing support for scaled clustering by transforming a local cluster address into a corresponding global cluster address for each packet in an outbound path from any of the cluster nodes and by transforming a global cluster address into a corresponding local cluster address for each packet in an inbound path to any of the cluster nodes, wherein intra-node cluster addressing is transparent to inter-node cluster address changes, and wherein re-configuration of the scalable clustered system requires no address re-assignments yet allowing the end nodes in the cluster nodes to maintain connectivity therebetween.

2. A scalable clustered system as in claim 1 wherein the local fabric and global fabric provide local and global clustering support infrastructures, respectively, and wherein global routing and global fabric topology are transparent to end nodes in operation.

3. A scalable clustered system as in claim 1 wherein the local and global fabric are each configured with one or more routers and/or switches.

4. A scalable clustered system as in claim 1 wherein each of the end nodes is an addressable device representing a resource such as a microprocessor, a central processing unit (CPU), a memory, an input/output (I/O) device controller or a server.

5. A scalable clustered system as in claim 1, the re-configuration of which, by addition thereto or removal therefrom of any cluster node, is possible while maintaining the configuration of any of the other cluster nodes, and wherein configuration of any of the cluster nodes is possible without reconfiguration of any of the other cluster nodes.

6. A scalable clustered system as in claim 1 wherein connections in the local and global fabric can be modified while maintaining correct packet transmission behavior.

7. A scalable clustered system as in claim 1 wherein the number of end nodes in each cluster node corresponds to an attribute of address fields in packets.

8. A scalable clustered system as in claim 1 wherein the NNA is configured to prevent a duplicated assignment of any end node address in establishing end node membership in a cluster node.

9. A scalable clustered system as in claim 1 wherein the scalable clustered system has a hierarchical topology and wherein each of the cluster nodes has either a flat or a hierarchical topology.

10. A scalable clustered system as in claim 1 wherein local traffic of packets within any of the cluster nodes is not routed to its respective NNA.

11. A scalable clustered system as in claim 1 wherein global traffic of packets to and from each of the cluster nodes is routed via its respective NNA.

12. A scalable clustered system as in claim 1 wherein the NNA includes a mask register for transforming the global/local cluster addresses in which bit substitutions can be made before an entire address has arrived at the NNA.

13. A scalable clustered system as in claim 1 having hierarchical topology and address-identification scheme that relative to flat non-hierarchical topology require smaller address fields and routing tables.

14. A scalable clustered system as in claim 1 wherein the NNA is a semiconductor chip.

15. A scalable clustered system as in claim 1 wherein the NNA has symmetrically built ends with one end being connected to the local fabric and the other end being connected to the global fabric and wherein a capability replicated at each end includes status indications and control enable bits.

5

16. A scalable clustered system as in claim 15 wherein the control enable bits include a shutdown on missing clock enable, a replace source address enable, a replace destination address enable, a destination address checking enable and a pass-through enable.

10

17. A scalable clustered system as in claim 1 wherein the NNA has a mode control register the contents of which determines which mode of operation the NNA assumes.

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

18. A scalable clustered system as in claim 17 wherein the NNA modes of operation include a pass-through mode, a conversion mode, a clock-check mode, an error check mode an error recovery mode and a shutdown mode.

19. A scalable clustered system as in claim 1 wherein, while in the conversion mode, the NNA is configured to perform source or destination cluster identification (ID) translation, clock checking and destination address checking.

20. A scalable clustered system as in claim 18 wherein the pass-through mode is a default mode of the NNA upon initialization.

21. A scalable clustered system as in claim 1 wherein NNA includes data replacement registers programmable with information for converting local cluster address to global cluster address and global cluster address to local cluster address.

22. A method for scaling a clustered system, comprising:
operatively linking two or more cluster nodes via a global fabric in order to form a larger clustered system, each of the cluster nodes having end nodes and a local fabric interconnecting the end nodes;

routing global packet traffic between the two or more cluster nodes in the larger clustered system via the global fabric;

routing local packet traffic between the one or more end nodes within each of the cluster nodes via the local fabric; and

5 operatively interposing an NNA between the local fabric and the global fabric, the NNA being configured as a fully symmetrical translation device, the NNA providing support for scaled clustering by transforming a local cluster identification (ID) into a corresponding global cluster ID for each packet in an outbound path from any of the cluster nodes and by transforming a global cluster ID into a corresponding local cluster ID for each packet in an inbound path to any
10 of the cluster nodes, wherein intra-node cluster addressing is transparent to inter-node cluster address changes, and wherein re-configuration of the scalable clustered system requires no ID re-assignments yet allowing the end nodes in the cluster nodes to maintain connectivity therebetween.

15 23. A method as in claim 22 wherein in the outbound path the NNA performs steps including forwarding control and status symbols,
 replacing a local cluster ID value with a global cluster ID value in a packet source address field if the NNA operates in a conversion mode, and
 establishing a cyclic redundancy check (CRC) value in an outbound packet CRC
20 field based on whether a correct or incorrect CRC value is detected.

25 24. A method as in claim 23 wherein the CRC value is re-computed if a correct CRC value is detected, a destination cluster ID is enabled and there is no destination cluster ID mismatch.

30 25. A method as in claim 24 wherein in the inbound path, the NNA performs steps including forwarding control and status symbols,
 verifying proper routing by checking a destination cluster ID field in an inbound packet,
 detecting a correct or incorrect CRC value in the inbound packet,
 setting a status bit if the incorrect CRC value is detected, and

replacing a global cluster ID value with a local cluster ID value in a packet destination address field if the NNA operates in a conversion mode.

26. In a scalable clustered system, a computer readable medium embodying computer program code configured to cause that system to perform steps for configuring and scaling that system, comprising:

operatively linking two or more cluster nodes via a global fabric in order to form a larger clustered system, each of the cluster nodes having end nodes and a local fabric interconnecting the end nodes;

routing global packet traffic between the two or more cluster nodes in the larger clustered system via the global fabric;

routing local packet traffic between the one or more end nodes within each of the cluster nodes via the local fabric; and

operatively interposing an NNA between the local fabric and the global fabric, the NNA being configured as a fully symmetrical translation device, the NNA providing support for scaled clustering by transforming a local cluster address into a corresponding global cluster address for each packet in an outbound path from any of the cluster nodes and by transforming a global cluster address into a corresponding local cluster address for each packet in an inbound path to any of the cluster nodes, wherein intra-node cluster addressing is transparent to inter-node cluster address changes, and wherein re-configuration of the scalable clustered system requires no address re-assignments yet allowing the end nodes in the cluster nodes to maintain connectivity therebetween.

27. A computer readable medium as in claim 26 wherein the computer program code is further configured to cause the NNA to prevent

assignment of a same local address to two end nodes in the same cluster node, and associating an end node with more than one cluster node.

28. A scalable clustered system, comprising:

global interconnection means for operatively linking two or more cluster nodes in order to form a larger clustered system, each of the cluster nodes having end nodes;

local interconnection means for operatively linking the end nodes within each of the cluster nodes;

global routing means for routing global packet traffic between the two or more cluster nodes in the larger clustered system via the global interconnection means;

5 local routing means for routing local packet traffic between the one or more end nodes within each of the cluster nodes via the local fabric means; and

translation means operatively interposed between the local interconnection means and the global interconnection means, the translation means being configured as a fully symmetrical translation device, the translation means providing support for scaled clustering by transforming
10 a local cluster address into a corresponding global cluster address for each packet in an outbound path from any of the cluster nodes and by transforming a global cluster address into a corresponding local cluster address for each packet in an inbound path to any of the cluster nodes, wherein intra-node cluster addressing is transparent to inter-node cluster address changes, and wherein re-configuration of the scalable clustered system requires no address re-assignments
15 yet allowing the end nodes in the cluster nodes to maintain connectivity therebetween.

29. A scalable super-clustered system, comprising:

a primary level of hierarchy configured with a plurality of global fabrics each of which interconnecting a plurality of cluster nodes to form one or more primary-level clusters, each cluster node including a node-level node naming agent (NNA), a local fabric and one or more
20 end nodes interconnected via the local fabric, the node-level NNA being configured as a fully symmetrical translation device interposed between its local fabric and one of the global fabrics to which the cluster node is connected; and

an upper level of the hierarchy configured with one or more upper-level global fabrics
25 each of which interconnecting a plurality of the primary-level clusters to form one or more upper-level clusters, each primary-level cluster including a primary-level NNA that is configured as a fully symmetrical translation device interposed between its global fabric and a particular one of the upper-level global fabrics to which the primary-level cluster is connected,

wherein intra-node cluster addressing is transparent to inter-node cluster address changes
30 at any level of the hierarchy, and wherein re-configuration of the scalable super-clustered system

requires no address re-assignments yet allowing the end nodes in the cluster nodes to maintain connectivity therebetween.

30. A scalable super-clustered system as in claim 29, wherein the node-level NNA is configured to provide support for super-scaled clustering by transforming a local primary-level cluster address into a corresponding upper-level global cluster address for each packet in an outbound path from any of the primary-level clusters and by transforming an upper-level global cluster address into a corresponding local primary-level cluster address for each packet in an inbound path to any of the primary-level clusters.

31. A scalable super-clustered system as in claim 29, wherein the primary-level NNA is configured to provide support for super-scaled clustering by transforming a local cluster address into a corresponding global cluster address for each packet in an outbound path from any of the cluster nodes and by transforming a global cluster address into a corresponding local cluster address for each packet in an inbound path to any of the cluster nodes.

32. A scalable super-clustered system as in claim 29, wherein the NNA is configured with a software algorithm to prevent a duplicate assignment of any address so that, at any point in time, an end node cannot be a member of more than one of the plurality of cluster nodes.

33. A scalable super-clustered system as in claim 29 capable of operating in a TCP-IP (transmission control protocol – Internet protocol) network environment.

34. A scalable clustered system as in claim 1, wherein the NNA is configured with a software algorithm to prevent a duplicate assignment of any address so that, at any point in time, an end node cannot be a member of more than one of the plurality of cluster nodes.

35. A scalable clustered system as in claim 1 capable of operating in a TCP-IP network environment.

36. A scalable clustered system as in claim 28 capable of operating in a TCP-IP (transmission control protocol – internet protocol) network environment.

37. A method as in claim 22 implemented in a TCP-IP (transmission control protocol – internet protocol) network environment.

38. A scalable super-clustered system as in claim 29 further comprising:
higher levels of the hierarchy each of which including a higher-level fabric interconnecting clusters of the level below, each such cluster including a sub-cluster NNA interposed between that cluster and the higher-level fabric to which it is connected.

39. A scalable super-clustered system as in claim 29, wherein the local cluster address is a fixed cluster number of a cluster node, and wherein the global cluster address is a cluster number assigned to the cluster node during cluster configuration.

40. A scalable clustered system as in claim 1, wherein the local cluster address is a fixed cluster number of a cluster node, and wherein the global cluster address is a cluster number assigned to the cluster node during cluster configuration.